DOCUMENT RESUME

ED 464 132 TM 033 815

AUTHOR	Raju, Nambury S.; Arenson, Ethan
TITLE	Developing a Common Metric in Item Response Theory: An Area-Minimization Approach.
PUB DATE	2002-04-00
NOTE	25p.; Paper presented at the Annual Meeting of the National Council on Measurement in Education (New Orleans, LA, April 2-4, 2002).
PUB TYPE	Reports - Descriptive (141) -- Speeches/Meeting Papers (150)
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	*Item Response Theory; *Test Items
IDENTIFIERS	Calibration; *Linking Metrics

ABSTRACT

An alternative method of finding a common metric for separate calibrations through the use of a common (anchor) set of items is presented. Based on Raju's (1988) method of calculating the area between the two item response functions, this (area-minimization) method minimizes the sum of the squared exact unsigned areas of each of the common items. This new method and five other currently available linking methods are illustrated with an empirical example. For this purpose, data from a calibration administration of two forms of a statewide high school algebra test were used. The need for additional research in this area, especially to establish the degree of congruence among various linking methods, is strongly recommended. (Contains 5 tables and 18 references.) (Author/SLD)

RUNNING HEAD: Common Metric in IRT

Developing a Common Metric in Item Response Theory: An Area-Minimization Approach

Nambury S. Raju

Illinois Institute of Technology

Ethan Arenson

CTB/McGraw-Hill

Paper presented at the 2002 Annual Meeting of the National Council on Measurement in Education, New Orleans, LA. The authors would like to express their appreciation to Dr. T. C. Oshima for her assistance with the IPLINK computer program.

Abstract

An alternative method of finding a common metric for separate calibrations through the use of a common (anchor) set of items is presented. Based on Raju's (1988) method of calculating the area between two item response functions, this (area-minimization) method minimizes the sum of the squared exact unsigned areas of each of the common items. This new method and five other currently available linking methods are illustrated with an empirical example. The need for additional research in this area, especially to establish the degree of congruence among various linking methods, is strongly recommended.

Developing a Common Metric in Item Response Theory: An Area-Minimization Approach

It is well known in item response theory (IRT) that estimates of item parameters and abilities from two different calibrations are not directly comparable because they may not be on a common metric (Hambleton, Swaminathan, & Rogers, 1991; Lord, 1980). In general, two independent IRT calibrations result in two different metrics for item parameters and examinee abilities. In many applications of IRT, such as vertical and horizontal equating and differential item functioning (DIF), estimates of item parameters and abilities from two independent calibrations must be placed on a common metric prior to their use. The development of such a common metric involves transforming the metric from one calibration to the metric from the other calibration with two appropriately defined linking constants: A multiplicative constant (A) and an additive constant (B). For example, an IRT calibration of a test with the 2-parameter logistic (2-PL) model will result in an a-parameter and a b-parameter for each item in the test. Let the item parameter estimates from the first calibration be denoted as a_{i1} and b_{i1} for item i. Similarly, let the item parameter estimates for item i from the second calibration be denoted as a_{i2} and b_{i2}. The transformation that puts the metric from the second calibration on to the metric from the first calibration may be defined as (Hambleton et al., 1991; Lord, 1980):

$$a_{i2}^{*} \equiv \frac{a_{i2}}{A}, \qquad (1)$$

and

$$b_{i2}^{*} \equiv Ab_{i2} + B. \qquad (2)$$

The transformed item parameters (a^*_{i2} and b^*_{i2}) from the second calibration are then said be on the same metric as the item parameters (a_{i1} and b_{i1}) from the first calibration. The transformation given in Equation (2) is also valid for transforming the ability (θ) estimates from the second calibration to the metric associated with the first calibration.

Current Methods for Developing a Common Metric

There are currently several methods for deriving the two transformation/linking constants (A and B). Five of these methods are briefly described below.

Mean-Mean Method

The Mean-Mean method is a variant of the method developed by Loyd and Hoover (1980). It sets the means of the a- and b-parameters of the common items in the second test equal to those of the first test. That is, the transformation constants are defined as follows:

$$A = \frac{M_{a2}}{M_{a1}}, \tag{3}$$

and

$$B = M_{b1} - AM_{b2}, \tag{4}$$

where M_{ai} and M_{bi} refer to the means of a-parameters and b-parameters, respectively, for test/calibration i or for the set of common items from test/calibration i. Additional information about this method may be found in Baker and Al-Karni (1991) and Kolen and Brennan (1995).

Mean-Sigma Method

In the Mean-Sigma method (Marco, 1977), the means and standard deviations of the b-parameters from the first and second administrations/calibrations determine the A and B constants:

$$A = \frac{SD_{b(1)}}{SD_{b(2)}} \tag{5}$$

and

$$B = M_{b(1)} - AM_{b(2)}, \tag{6}$$

where $SD_{b(i)}$ refers to the standard deviation of b-parameters in test calibration i or in the set of common items from test/calibration i.

χ^2-Method

The χ^2-Method (Divgi, 1985) uses the standard errors of the item parameter estimates from both calibrations. Define $\mathbf{\Sigma}$ such that for item i,

$$\Sigma_i = \begin{bmatrix} I_{i,aa} & I_{i,ab} \\ I_{i,ba} & I_{i,bb} \end{bmatrix}^{-1}, \tag{7}$$

where, for the two-parameter model, $I_{i,aa}$, $I_{i,ab}$, and $I_{i,bb}$ are defined by Lord (1980, p. 191) as:

$$I_{i,aa} = D^2 \sum_{a=1}^{N} \left[(\theta_a - b_i)^2 P_{ia} (1 - P_{ia}) \right], \tag{8}$$

$$I_{i,ab} = -a_i D^2 \sum_{a=1}^{N} \left[(\theta_a - b_i) P_{ia} (1 - P_{ia}) \right], \tag{9}$$

and

$$I_{i,bb} = (Da_i)^2 \sum_{a=1}^{N} \left[P_{ia} (1 - P_{ia}) \right]. \tag{10}$$

It can be shown that for the equated item parameters from the second calibration,

$$I^*_{i,aa} = \frac{I_{i,aa}}{A^2}, \tag{11}$$

$$I^*_{i,ab} = I_{i,ab}, \tag{12}$$

and

$$I^*_{i,bb} = A^2 I_{i,bb}. \tag{13}$$

To distinguish among the variance-covariance matrices of the parameter estimates for item i from the first and second calibrations, let $\mathbf{\Sigma}_{i1}$ and $\mathbf{\Sigma}^*_{2,i}$ denote, respectively, the item parameter variance-covariance matrices from the first calibration and the second calibration after equating. Letting

$$\mathbf{\Delta} = \begin{bmatrix} a_{i1} - a^*_{i2} \\ b_{i1} - b^*_{i2} \end{bmatrix}, \tag{14}$$

the quadratic form to be minimized is

$$Q = \mathbf{\Delta}' \left(\mathbf{\Sigma}_{1i} + \mathbf{\Sigma}^*_{2i} \right)^{-1} \mathbf{\Delta}. \tag{15}$$

In this investigation, a variant of Equation (15) is used based on methods proposed by Oshima et al. (2000). Specifically, for an n-item test, the function to be minimized is

$$Q^* = \frac{1}{2n} \left(\sum_{i=1}^{n} \left(a_{i1} - a^*_{i2} \right)^2 + \sum_{i=1}^{n} \left(b_{i1} - b^*_{i2} \right)^2 \right). \tag{16}$$

<ins>Test Characteristic Curve (TCC) Method</ins>

The Test Characteristic Curve (TCC) method (Stocking & Lord, 1983) uses item parameter estimates from both calibrations, as well as a spaced set of abilities, in minimizing the appropriate multivariate functions needed for estimating the linking constants (A and B). Following Oshima et al. (2000), the function to be minimized in the TCC approach may be expressed as:

$$f_1(A,B)=\frac{1}{N}\sum_{\theta}\left[\sum_{i=1}^{n}P_{1i}(\theta)-\sum_{i=1}^{n}P_{2i}^{*}(\theta)\right]^2 \tag{17}$$

where N stands for the number of spaced abilities used and n represents the number of items. The $P_{1i}(\theta)$ and $P_{2i}^{*}(\theta)$, for the 2-PL model, may be expressed as:

$$P_{1i}(\theta)=\frac{e^{Da_{1i}(\theta-b_{1i})}}{1+e^{Da_{1i}(\theta-b_{1i})}}, \tag{18}$$

and

$$P_{2i}^{*}(\theta)=\frac{e^{Da_{2i}^{*}(\theta-b_{2i}^{*})}}{1+e^{Da_{2i}^{*}(\theta-b_{2i}^{*})}}; \tag{19}$$

where D and e are constants (Hambleton, Swaminathan, & Rogers, 1991). The general idea in the TCC approach is to find A and B which minimize Equation (17).

Item Characteristic Curve Method

Like the TCC method, the Item Characteristic Curve (ICC) Method (Haebara, 1980) uses item parameter estimates from both calibrations, as well as a spaced set of abilities to minimize the multivariate function

$$f_1(A,B)=\frac{1}{N}\sum_{\theta}\sum_{i=1}^{n}\left[P_{1i}(\theta)-P_{2i}^{*}(\theta)\right]^2. \tag{20}$$

Some of these approaches, originally developed for unidimensional, dichotomous IRT models, were later expanded to include multidimensional, dichotomous IRT models

(Davey, Oshima, & Lee, 1996; Oshima, Davey, & Lee, 2000) and unidimensional, graded response IRT models (Baker, 1992). Computer programs for estimating the needed transformation constants are also available from Baker (1992) and Lee and Oshima (1996). Additional information about the previously described linking methods may be found in Kolen and Brennan (1995).

A careful examination of Equations (17) and (20) leads to two important observations: First, both equations depend on the specific number and type of thetas chosen for minimization. Stocking and Lord (1983) recommended 200 spaced thetas, and Baker and Al-Karni's (1991) example used 21 spaced thetas. As Oshima, Davey, and Lee (2000) noted, there is a certain degree of arbitrariness in the number and type of thetas chosen. Also, Kolen and Brennan (1995) describe 5 different examinee or theta selection procedures for use in developing a common metric. Second, in the TCC approach, there is the potential for cancellation at the test level; that is, positive and negative differences in item probabilities, $P_{1i}(\theta) - P^{*}_{2i}(\theta)$, may cancel each other out. However, this is not a concern for the ICC approach because the item level probability differences are squared prior to summing them across thetas.

The purpose of this investigation is to offer a new approach for developing a common IRT metric that avoids the problems of arbitrariness (due to the number and type of thetas used) and cancellation effects. This new approach is based on Raju's (1988) exact unsigned area measure. Also included in this presentation is an empirical example to illustrate the new procedure as well as the five previously described linking procedures.

An Area-Minimization Approach

According to Raju (1988), the exact unsigned area (EUA) between two ICCs for item i, based on the 2-PL model, may be expressed as

$$EUA_i = |H_i|, \tag{21}$$

where

$$H_i = \frac{2(a_{1i} - a_{2i}^{\bullet})}{Da_{1i}a_{2i}^{\bullet}} \ln\left\{1 + \exp\left[\frac{Da_{1i}a_{2i}^{\bullet}(b_{1i} - b_{2i}^{\bullet})}{a_{1i} - a_{2i}^{\bullet}}\right]\right\} - (b_{1i} - b_{2i}^{\bullet}) \tag{22}$$

and, as before, subscripts 1 and 2 refer to the first and second calibrations, respectively. Substituting a_{i2} and b_{i2} from Equations (1) and (2) into Equation (22), one obtains

$$H_i = \frac{2(Aa_{1i} - a_{2i})}{Da_{1i}a_{2i}} \ln\left\{1 + \exp\left[\frac{Da_{1i}a_{2i}}{Aa_{1i} - a_{2i}}\left(b_{1i} - Ab_{2i} - B\right)\right]\right\} - \left(b_{1i} - Ab_{2i} - B\right). \tag{23}$$

Note that Equation (23) is independent of theta. Therefore, the potential problems from the TCC and ICC methods, Equations (17) and (20) respectively, are not a concern for the area-minimization approach.

Solution for A and B

In view of Equation (21), the function to be minimized across n items in solving for A and B may be expressed as:

$$f_3(A,B) = |H_1| + |H_2| + \ldots + |H_n|. \tag{24}$$

The absolute values in Equation (24) may not be easy to handle (mathematically), so an alternative but equivalent expression for minimization may be written as

$$f_4(A,B) = H_1^2 + H_2^2 + \ldots + H_n^2 \ . \tag{25}$$

Since Equations (24) and (25) contain either absolute or squared quantities, there is no potential for cancellation effects in the area-minimization approach. The partial derivatives of $f_4(A,B)$ with respect to A and B can be written as:

$$\frac{\partial f_4(A,B)}{\partial A} = 2H_1 \frac{\partial H_1}{\partial A} + 2H_2 \frac{\partial H_2}{\partial A} + \ldots + 2H_n \frac{\partial H_n}{\partial A} \tag{26}$$

and

$$\frac{\partial f_4(A,B)}{\partial B} = 2H_1 \frac{\partial H_1}{\partial B} + 2H_2 \frac{\partial H_2}{\partial B} + \ldots + 2H_n \frac{\partial H_n}{\partial B}. \tag{27}$$

Using Equation (23), the partial derivatives of H_i may be expressed, after simplification, as:

$$\frac{\partial H_i}{\partial A} = \frac{2}{Da_{2i}} \ln\left[1+\exp(Y)\right] - \frac{2\exp(Y)}{1+\exp(Y)} \left[\frac{a_{1i}(b_{1i}-B) - a_{2i}b_{2i}}{Aa_{1i}-a_{2i}}\right] + b_{2i} \tag{28}$$

and

$$\frac{\partial H_i}{\partial B} = \frac{-2\exp(Y)}{1+\exp(Y)} + 1, \tag{29}$$

where

$$Y = \frac{Da_{1i}a_{2i}}{Aa_{1i}-a_{2i}}(b_{1i} - Ab_{2i} - B). \tag{30}$$

Given the mathematical complexity of Equations (28) and (29), it will not be easy to find the linking constants (A and B) that minimize this equation. However, the Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm (Press, Flannery, Teukolsky, & Vetterling, 1986) can be used to solve for A and B, using Equations (28) through (30). The BFGS algorithm improves upon the Davidon-Fletcher-Powell algorithm, which is commonly used in psychometric practice (e.g., for Stocking and Lord equating).

The above procedure for estimating *A* and *B* within the two-parameter logistic model is equally valid for the three-parameter model, provided each item has the same *c* parameter in both calibrations. One way to obtain such a common *c* parameter is to set the estimates of *c* from one calibration to equal those from the other (Stocking & Lord, 1983). According to Divgi (1985), the estimation of *c* parameters is not affected by the change of metric and, hence, may be ignored in developing a common metric.

With respect to the Rasch model, where only one linking constant (*B*) is needed for metric transformation, the area minimization approach results in a solution for *B*, which is equal to the difference in the means of *b* parameters. That is,

$$B = M_{b1} - AM_{b2}, \tag{31}$$

Additional details and the proof of this result for the Rasch model are given in Arenson and Raju (2002).

An Illustration and Discussion

To illustrate the area-minimization approach, as well as other approaches, for developing a common metric, data from a calibration administration (towards the end of the Spring 2001 semester) of two forms (Forms K and L) of a statewide high school algebra test were used. The test constitutes a partial requirement for graduation from high school. Both forms, containing 55 items each, were calibrated concurrently. The distributions of raw scores were similar for both forms, as shown in Table 1. Raw scores were converted to a scale with a mean of 500 and a standard deviation of 50. Although the distributions of raw scores were similar, the tests contained different items, except for the 13 anchor items. Item parameters for these 13 items were used to transform the metric underlying Form L onto the metric underlying Form K. Six different transformations were obtained, one for each of the six linking methods described above.

The two-parameter logistic model (Lord, 1980) was used to estimate the item parameters, separately, for Forms K and L. The two forms were calibrated with PARDUX (Burket, 1991). Once the item parameters for the 13 anchor items were known, separate linking constants (*A* and *B*) for each of the six methods were obtained and are reported in Table 2. The linking constants for the Mean-Mean and Mean-Sigma methods were computed with a program written in SAS. The linking constants for Haebera's ICC method and Divgi's chi-square were obtained with the IPLINK program (Lee & Oshima, 1996). It should be noted that linking constants for Divgi's method were based on a simplified version of Equation (15) (Oshima et al., 2000). The linking constants for Stocking and Lord's TCC method were obtained with the PARDUX program. Finally, the linking constants associated with the area-minimization approach were obtained with a computer program specially written for this research. Means and standard deviations of the *a*- and *b*-parameters for the anchor items in Form K are shown in Table 3. Also shown in this table are the means and standard deviations of the one unequated and 6 equated item parameters for Form L. Table 4 displays the *a*-parameter estimates for the common items from Form K, as well as the unequated parameter estimates from form L. In addition, Table 4 shows the equated estimates for each of the methods described. Table 5 displays similar information for the *b*-parameters.

The equating constants (see Table 2) appear to be quite similar across the 6 linking procedures. The similarities between the Form K and the equated Form L parameter estimates are best captured in Figures 1 and 2. The *a*-parameters are close to 1 and the b-parameters are close to 0 across the five procedures. This result is probably not too surprising in view of the fact that the two forms were quite comparable in terms of their raw score means and standard deviations. This is not to say that the estimated linking constants will be this similar for other tests and/or forms or for the same test with different sub-populations as in the case of DIF research. The current example is designed simply to illustrate the area-minimization approach, while presenting comparable data from the other available linking procedures. This example is not intended to offer an evaluation of the various procedures. There is certainly a need for a comprehensive assessment of the various linking procedures, hopefully with recommendations for practitioners. Oshima et al. (2000) recently reported a Monte Carlo assessment of different linking procedures in the multi-dimensional IRT context. A similar study in the unidimensional IRT framework is highly desirable. Kolen and Brennan (1995) also recommend the need for assessing the comparability and accuracy of the known linking procedures. As Oshima et al. and Divgi (1985) noted, a major problem that one is likely to encounter in such an investigation is the question of what criterion to use for evaluating the results from different linking procedures. If the area measure is used as the criterion for assessing the accuracy of various linking procedures, the area-minimization method is likely to perform better than the other method. If the difference between two TCCs is used as the criterion, the Stocking and Lord's method may do better than the other methods. So, there is a definite need for defining an appropriate (or impartial) criterion for assessing the accuracy of various linking procedures. Finally, there is a need for extending the area-minimization approach to the polytomous case.

References

Arenson, E. A., & Raju, N. S. (April, 2002). An Area-Minimization Approach for the Equating of Test Scores with the Rasch Model. Paper presented at the annual meeting of the Institute for Objective Measurement, New Orleans, LA.

Baker, F.B. (1992). Equating tests under the graded response model. Applied Psychological Measurement, 16, 87-96.

Baker, F. B. (1995). EQUATE 2.1: Computer program for equating two metrics in item response theory [Computer Program]. Madison: University of Wisconsin, Laboratory of Experimental Design.

Baker, F.B., & Al-Karni, A. (1991). A comparison of two procedures for computing IRT equating constants. Journal of Educational Measurement, 28, 147-162.

Davey, T.C., Oshima, T.C., Lee, K. (1996). Linking multidimensional item calibrations. Applied Psychological Measurement, 20, 405-416.

Burket, G. (1991). PARDUX [Computer Program]. Monterey, CA: CTB/McGraw-Hill.

Divgi, D.R. (1985). A minimum chi-square method for developing a common metric in item response theory. Applied Psychological Measurement, 9, 413-415.

Haebara, T. (1980). Equating logistic ability scales by a weighted least squares method. Japanese Psychological Research, 22, 144-149.

Kolen, M.J., & Brennan, R.L. (1995). Test equating: Methods and practices. New York: Springer-Verlag.

Lee, K., & Oshima, T. C. (1996). IPLINK: Multidimensional and unidimensional item parameter linking in item response theory. Applied Psychological Measurement, 20, 230.

Loyd, B. H., & Hoover, H. D. (1980). Vertical equating using the Rasch Model. Journal of Educational Measurement, 17, 179-193.

Lord, F. M. (1980). Applications of item response theory to practical testing problems. Hillside, NJ: Erlbaum.

Marco, G. L. (1977). Item characteristic curve solutions to three intractable testing problems. Journal of Educational Measurement, 14, 139-160.

Press, W.H., Flannery, B.P., Teukolsky, S.A., & Vetterling. (1986). Numerical Recipes: The art of scientific computing. Cambridge, England: Cambridge University Press.

Oshima, T.C., Davey, T.C., & Lee, K. (2000). Multidimensional linking: four practical approaches. Journal of Educational Measurement, 37, 357-373.

Raju, N. S. (1988). The area between two item characteristic curves. Psychometrika, 53, 495-502.

Raju, N. S. (1990). Determining the significance of estimated signed and unsigned areas between two item response functions. Applied Psychological Measurement, 14, 197-207.

Stocking, M.L., & Lord, F.M. (1983). Developing a common metric in item response theory. Applied Psychological Measurement, 7, 201-210.

Footnote

[1]The Loyd and Hoover (1980) method was designed for the Rasch model. They defined the multiplicative constant as $A \equiv \frac{a_2}{a_1}$, the ratio of the discriminant constant of the second calibration to that of the first calibration.

Table 1

Descriptive statistics for Forms K and L

		Raw Score		Scale Score		
Form	N	Mean	SD	Mean	SD	Reliability
K	*6994*	32.03	10.707	554.5	35.65	0.916
L	*6941*	32.97	10.388	551.0	36.90	0.916

Table 2

Equating Linking Constants

	Mean-Mean	Mean-Sigma	ICC (Haebera)	χ^2 (Divgi)	Stocking & Lord	Area-Min.
A	0.9758	0.9975	0.9792	0.9752	0.9761	0.9852
B	0.1180	0.1019	0.1146	0.1035	0.1020	0.1123

Table 3

Summary Statistics of Unequated and Equated Item Parameters

	a		*b*	
	Mean	SD	Mean	SD
Form K	1.205	0.319	0.842	0.771
Form L				
Unequated	1.176	0.298	0.742	0.773
Mean-Mean	1.235	0.328	0.941	0.752
Mean-Sigma	1.179	0.299	0.842	0.771
ICC (Haebera)	1.201	0.304	0.841	0.757
χ^2 (Divgi)	1.266	0.336	1.021	0.733
Stocking & Lord	1.203	0.305	0.826	0.753
Area-Minimization	1.194	0.302	0.843	0.759

Table 4

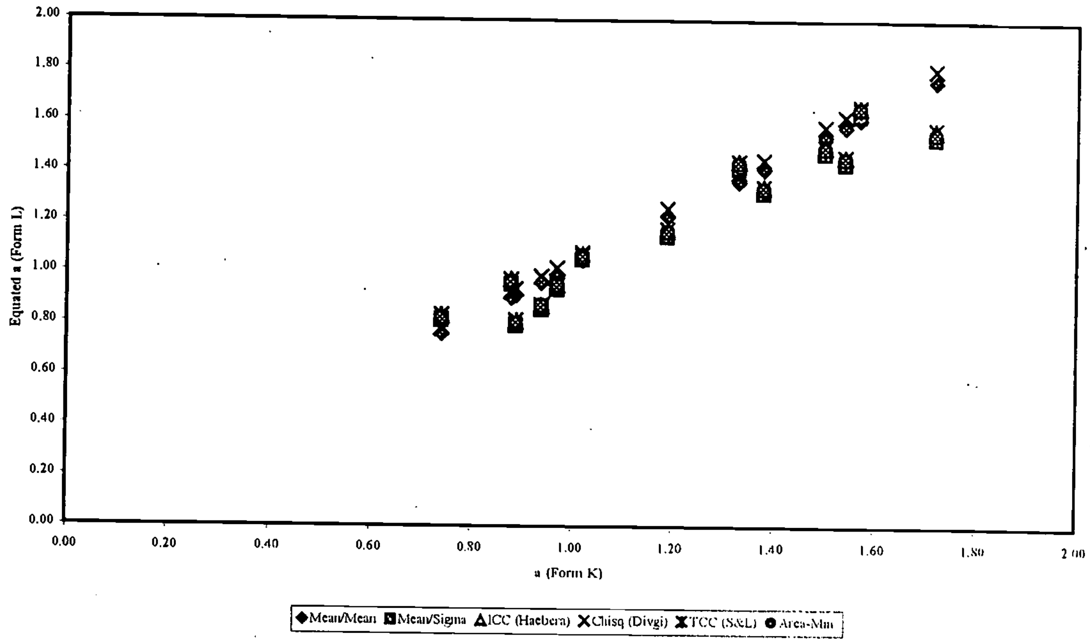
a-Parameter Estimates for Common Items in Forms K and L

		Form L						
Item No.	Form K	Unequated	Mean-Mean	Mean-Sigma	ICC (Haebera)	χ^2 (Divgi)	Stocking & Lord	Area-Min.
7	0.97	0.93	0.99	0.93	0.95	1.02	0.95	0.95
8	1.72	1.53	1.76	1.53	1.56	1.80	1.57	1.55
9	1.38	1.31	1.41	1.31	1.34	1.45	1.34	1.33
10	1.50	1.47	1.54	1.47	1.50	1.58	1.50	1.49
29	1.57	1.62	1.61	1.62	1.65	1.65	1.66	1.65
30	1.54	1.43	1.58	1.43	1.46	1.62	1.46	1.45
31	0.88	0.95	0.90	0.95	0.97	0.92	0.97	0.96
32	1.19	1.14	1.22	1.14	1.16	1.25	1.17	1.16
33	0.94	0.85	0.96	0.85	0.87	0.98	0.87	0.87
43	1.33	1.41	1.36	1.41	1.44	1.39	1.44	1.43
44	0.74	0.81	0.76	0.81	0.83	0.78	0.83	0.82
45	1.02	1.05	1.05	1.05	1.07	1.08	1.07	1.06
46	0.89	0.79	0.91	0.79	0.81	0.93	0.81	0.80

Table 5

b-parameter estimates for common items in Forms K and L

		Form L						
Item No.	Form K	Unequated	Mean-Mean	Mean-Sigma	ICC (Haebera)	χ^2 (Divgi)	Stocking & Lord	Area-Min.
7	-0.43	-0.60	-0.30	-0.50	-0.47	-0.19	-0.49	-0.48
8	-0.05	-0.22	0.07	-0.12	-0.10	0.17	-0.11	-0.10
9	0.68	0.62	0.78	0.72	0.72	0.86	0.71	0.72
10	0.09	-0.04	0.21	0.06	0.08	0.31	0.07	0.08
29	0.70	0.56	0.80	0.66	0.66	0.88	0.65	0.67
30	0.37	0.27	0.48	0.37	0.38	0.57	0.37	0.38
31	0.87	0.82	0.97	0.92	0.92	1.05	0.90	0.92
32	1.49	1.40	1.57	1.50	1.49	1.63	1.47	1.49
33	1.59	1.62	1.67	1.72	1.70	1.73	1.68	1.70
43	0.86	0.79	0.96	0.89	0.89	1.04	0.87	0.89
44	2.46	2.15	2.52	2.25	2.22	2.56	2.20	2.23
45	0.96	0.99	1.05	1.09	1.08	1.13	1.06	1.08
46	1.36	1.29	1.45	1.39	1.38	1.52	1.36	1.38


2.00
1.80
1.60
1.40
1.20
1.00
0.80
0.60
0.40
0.20
0.00
Equated a (Form L)
0.00
0.20
0.40
0.60
0.80
1.00
1.20
1.40
1.60
1.80
2.00
a (Form K)
Mean/Mean
Mean/Sigma
ICC (Haebera)
Chisq (Divgi)
TCC (S&L)
Area-Min

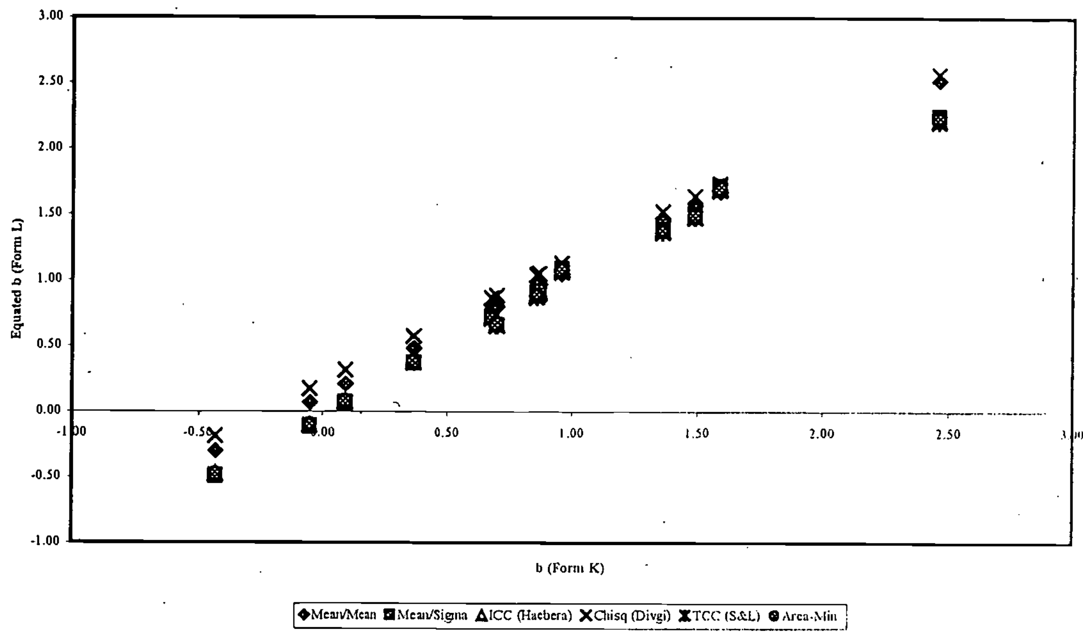

3.00
2.50
2.00
1.50
1.00
0.50
0.00
-0.50
-1.00
Equated b (Form L)
-1.00
-0.50
0.00
0.50
1.00
1.50
2.00
2.50
3.00
b (Form K)
Mean/Mean
Mean/Sigma
ICC (Haebera)
Chisq (Divgi)
TCC (S&L)
Area-Min

Figure Captions

Figure 1. Form K and Equated Form L a-Parameters

Figure 2. Form K and Equated Form L b-Prameters

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

ERIC®

REPRODUCTION RELEASE

TM033815

(Specific Document)

I. DOCUMENT IDENTIFICATION:

Title: Developing a Common Metric in Item Response Theory: An Area-Minimization Approach	
Author(s): Nambury S. Raju & Ethan A Arenson	
Corporate Source: CTB/McGraw Hill	Publication Date: 4/3/2002

II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, *Resources in Education* (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic media, and sold through the ERIC Document Reproduction Service (EDRS). Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce and disseminate the identified document, please CHECK ONE of the following three options and sign at the bottom of the page.

The sample sticker shown below will be affixed to all Level 1 documents	The sample sticker shown below will be affixed to all Level 2A documents	The sample sticker shown below will be affixed to all Level 2B documents
PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY Sample TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) 1	PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE, AND IN ELECTRONIC MEDIA FOR ERIC COLLECTION SUBSCRIBERS ONLY, HAS BEEN GRANTED BY Sample TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) 2A	PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY Sample TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) 2B
Level 1	Level 2A	Level 2B
☒	☐	☐
Check here for Level 1 release, permitting reproduction and dissemination in microfiche or other ERIC archival media (e.g., electronic) *and* paper copy.	Check here for Level 2A release, permitting reproduction and dissemination in microfiche and in electronic media for ERIC archival collection subscribers only	Check here for Level 2B release, permitting reproduction and dissemination in microfiche only

Documents will be processed as indicated provided reproduction quality permits.
If permission to reproduce is granted, but no box is checked, documents will be processed at Level 1.

I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce and disseminate this document as indicated above. Reproduction from the ERIC microfiche or electronic media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries.

Sign here, → please

Signature: Ethan A	Printed Name/Position/Title: Ethan Arenson / Assoc. Psychometrician	
Organization/Address: CTB/McGraw-Hill — Research, 20 Ryan Ranch Rd, Monterey CA 93940	Telephone: 831 393 7017	FAX:
	E-Mail Address: earenson@ctb.com	Date: 4/3/2002

(over)

III. DOCUMENT AVAILABILITY INFORMATION (FROM NON-ERIC SOURCE):

If permission to reproduce is not granted to ERIC, *or*, if you wish ERIC to cite the availability of the document from another source, please provide the following information regarding the availability of the document. (ERIC will not announce a document unless it is publicly available, and a dependable source can be specified. Contributors should also be aware that ERIC selection criteria are significantly more stringent for documents that cannot be made available through EDRS.)

Publisher/Distributor:
Address:
Price:

IV. REFERRAL OF ERIC TO COPYRIGHT/REPRODUCTION RIGHTS HOLDER:

If the right to grant this reproduction release is held by someone other than the addressee, please provide the appropriate name and address:

Name:
Address:

V. WHERE TO SEND THIS FORM:

Send this form to the following ERIC Clearinghouse:

University of Maryland
ERIC Clearinghouse on Assessment and Evaluation
1129 Shriver Laboratory
College Park, MD 20742
Attn: Acquisitions

However, if solicited by the ERIC Facility, or if making an unsolicited contribution to ERIC, return this form (and the document being contributed) to:

ERIC Processing and Reference Facility
4483-A Forbes Boulevard
Lanham, Maryland 20706

Telephone: 301-552-4200
Toll Free: 800-799-3742
FAX: 301-552-4700
e-mail: ericfac@inet.ed.gov
WWW: http://ericfac.piccard.csc.com

EFF-088 (Rev. 2/2000)